DOCUMENT RESUME

ED 057 075 TM 000 912

AUTHOR Lunneborg, Clifford E.
TITLE Adjusting Regression Weights for Criterion Group Similarity.
INSTITUTION Washington Univ., Seattle. Bureau of Testing.
PUB DATE Sep 71
NOTE 14p.; From symposium "People, Patterns and the Prediction of Academic Criteria," American Psychological Association Convention, Washington, D. C., September 1971

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Academic Performance; Achievement Tests; Aptitude Tests; *Bayesian Statistics; College Students; *Discriminant Analysis; *Experimental Groups; *Grade Point Average; Multiple Regression Analysis; Prediction; *Predictive Ability (Testing); Predictor Variables; Probability Theory; Symposia
IDENTIFIERS *Washington Pre College Testing Program

ABSTRACT

A Bayesian prediction strategy is outlined in which antecedent measures are divided into two subgroups. One subgroup is used to discriminate among criterion groups, the second to provide normal linear predictions for each group. Individualized regression constants are subsequently obtained by computing probabilities of group membership from the discriminating measures and weighting the group prediction equations by these probabilities. The technique is illustrated by the prediction of cumulative University of Washington GPA for student groups categorized by terminal university status using achievement and aptitude measures from the Washington Pre-College testing program. Errors on validation were slightly less for the adjusted predictions than for a single pooled prediction equation suggesting this may be a promising approach to the moderation of predictions. (Author)

N-AI
TM

Adjusting Regression Weights for Criterion Group Similarity[1]

Clifford E. Lunneborg

1

Bureau of Testing
University of Washington—Seattle

Bureau of Testing

University of Washington

September 1971

Adjusting Regression Weights for Criterion Group Similarity[1]

Clifford E. Lunneborg

A Bayesian prediction strategy is outlined in which antecedent measures are divided into two subgroups. One subgroup is used to discriminate among criterion groups, the second to provide normal linear predictions for each group. Individualized regression constants are subsequently obtained by computing probabilities of group membership from the discriminating measures and weighting the group prediction equations by these probabilities. The technique is illustrated by the prediction of cumulative University of Washington GPA for student groups categorized by terminal university status using achievement and aptitude measures from the Washington Pre-College testing program. Errors on validation were slightly less for the adjusted predictions than for a single pooled prediction equation suggesting this may be a promising approach to the moderation of predictions.

[1]Contribution to a symposium, "People, Patterns and the Prediction of Academic Criteria," American Psychological Association Convention, Washington, D. C., September 1971.

Bureau of Testing Project: 177

Adjusting Regression Weights for Criterion Group Similarity

Clifford E. Lunneborg

University of Washington

This symposium is testimony to the vigor with which psychometrists are currently pursuing productive alternatives to straightforward linear prediction of academic performance. This is not to say that the idea of patterns or profiles among predictors, interactions among predictors, differentially predictable subgroups, moderator variables and individualized predictions have not been with us for a fairly long time but it is in only recent years that, in application, there has se ed to be any appreciable probability of pay-off. This investigator, for one, could cite a series of personal forays against configural information over more than a decade, all of them halted with the grail still not in grasp. Yet the idea that certain kinds of antecedent information about students could function nonlinearly and interactively in predicting academic performance did not lose its appeal and might yet be rewarded.

How to mount a new attack? Heretofore this research had sought patterns among predictors and had bogged down in the empirical morass of the variety of kinds and numbers of patterns and their apparent instability. Where to look had been a problem--this pair or triad of measures might be as promising (or as disappointing) as the last one. A growing interest in a Bayesian approach to problem analysis suggested, in an informal way, that the interaction problem might be approached by breaking it up into stages and then reassembling the parts. In particular, the approach to be

explored was seen as something of a compromise between the use of differentially predictable groups and the classical moderator variable approach. It was predicated on the assumption that differentiable academic groups might be defined after the fact (say, by graduation status) rather than before. Normal linear prediction equations, potentially different for the several groups, could then be worked out. Since the group status of an entering freshman (for whom an academic success prediction is desired) would not be known no single one of these equations could be utilized. Rather, however, the full set of equations could be used, if the several equations were to be weighted or moderated consistent with the likelihood that the student would end up in the appropriate criterion group. Some second set of measures, observationally independent of those employed in the linear predictions, might be used through a discriminant function approach to provide these likelihood measures. This second set of measures would then function as moderator variables although they would likely not have a simple linear effect on the weights a[illegible] the variables in the prediction equations.

To phrase this strategy in terms of the prediction setting in which it was to be explored the following design was adopted: (1) A set of antecedent measures would be divided into two subsets; (2) With one subset, termed the predictor pool, normal linear prediction equations would be obtained for each of a number of previously defined criterion groups; (3) With the second subset of antecedent measures linear discriminant functions would be found which discriminated among these criterion groups; (4) For a new sample of students these discriminant functions would be used to estimate the relative probability of group membership across groups for each student; (5) These probabilities

would then be used to linearly weight the prediction equations from the several groups providing in essence an adjusted or individualized prediction equation for each validational student; (6) Finally, the accuracy of these predictions would be assessed relative to the accuracy of predictions from a single linear equation drawing upon all information in both sets of antecedent measures.

Procedure

Subjects. The groups of university students utilized in this preliminary study were six selected for an earlier discriminant function study (Lunneborg and Lunneborg, 1970) augmented by a seventh. This final group was chosen in an effort to bring the total sample of subjects into reasonably close correspondence to a complete entering freshman class. The original six groups consisted of those 1392 students who graduated in June, 1969, who had taken the Washington Pre-College (WPC) battery, and who had entered university directly from high school. On the basis of their academic major at the time of their graduation they had been classified into one of six areas: humanities (n = 285), physical science (n = 193), social science (n = 278), business (n = 123), biological science (n = 286), and engineering (n = 227). Together these six groups can be thought of as successful students. The seventh group, an unsuccessful one, was composed of those students who entered university directly from high school in the fall of 1965 who had also completed the WPC battery, and who were not in residence either in the fall of 1968 or the succeeding winter term. Oversimplifying, this seventh group (n = 388) consisted of those who had dropped by the wayside as far as the university was concerned sometime in the first three years following entrance.

For purposes of the analysis each of the seven groups was divided into two halves, by alternate assignment of subjects based on university serial number, providing weight development and validational subsamples. Each of these total subsamples, across the seven groups, contained 890 students. The smallest subgroup was made up of the 61 business graduates in the validation subsample and the largest subgroups were the two halves of the non-graduating sample, each with 194 students.

Measures. Antecedent measures available for each student included nine test scores from the WPC battery administered in the senior year of high school and six grade point averages (GPA's) based on the high school record. The nine tests were English usage, spelling, reading comprehension, vocabulary, mathematics achievement, applied mathematics, quantitative skills, mechanical reasoning, and spatial ability. The six high school GPA's were based on work through the junior year in English, foreign language, mathematics, natural science, social science, and electives courses. The single criterion measure was the cumulative GPA for all university work attempted.

The set of antecedent measures was subdivided into two sets. An earlier study (Lunneborg and Lunneborg, 1970) involving the six groups of graduating seniors established that those groups could be reliably discriminated by a pair of functions. Mathematics achievement, mechanical reasoning and vocabulary were strong determiners of the first and quantitative skills and spatial ability were most highly weighted in the second. These five measures, for the present study, were labelled the discriminant set of measures. The remaining four test scores (English usage, spelling,

reading comprehension, and applied mathematics) together with the six H.S. GPA's made up the potential predictor pool.

Analysis. Variance-covariance matrices and mean vectors were computed for the criterion and 10 potential predictors over the observations in the weight development sample. This was done separately for each of the seven subject groups. These results then formed the basis for a series of seven stepwise predictor selection and weightings utilizing Efroymson's (1960) approach. In each of these selections the maximum number of variables selected was arbitrarily limited to five. An eighth variance-covariance matrix and mean vector were computed across all seven weight development subsamples involving the five discriminant measures as well as the 10 measures for the predictor pool and the criterion. An unlimited stepwise selection of predictors of university GPA was obtained for these data. This part of the analysis yielded eight prediction equations for use with the validational sample: seven based on the criterion subgroups and involving only selections from the 10 predictors and one based on the full weighting sample and involving selections from these same 10 and from the five pre-identified discriminant measures.

Within group variance-covariance matrices and mean vectors, again for the seven weight development subsamples, were also computed for just the five discriminant measures. These results, together with sample sizes, were used in a modification of the BMD multiple groups discriminant analysis program (Dixon, 1968) to define a pair of linear functions which would maximally discriminate the seven groups. This completed the analysis of the weight development sample.

For the analysis of the validation sample a special program was prepared to compute for each member of that sample two predicted university GPA's. One of these, the total group prediction, was simply determined by the regression equation obtained from the free selection among the 15 antecedent measures. The second or adjusted weights prediction involved computing the seven subgroup predicted GPA's and then combining them linearly, weighting each of the seven by the estimated probability that the subject belonged to the subgroup for which a particular prediction was developed.

A major function of this special program, therefore, was the estimation of these probabilities of group membership. These estimates were computed from the calculated values of the two discriminant functions for each validational subject by the technique suggested by Cooley and Lohnes (1971, p. 267-268, Classification Rule III) for the case where the discriminant measures are assumed to be multivariate normal with differing variance-covariance matrices for the several unequal sized groups. In short, this technique converted scores on the five discriminant measures into a set of seven posterior probabilities reflecting the relative likelihood that the subject providing those scores would belong to each of the seven mutually exclusive criterion subgroups. These probabilities, scaled to sum to unity, were used to weight, for the individual, the seven subgroup predicted GPA's and provide the adjusted weights predicted GPA.

As a final function, the special program accumulated the squared discrepancies between the earned GPA and each of the two predictions, the total group prediction and the adjusted weights prediction, to permit an evaluation of the relative accuracy of the two sets.

Results

As in the earlier study not involving the unsuccessful criterion group (Lunneborg and Lunneborg, 1970) the two-dimensional discriminant space was spanned by a major function pitting mathematical achievement and mechanical reasoning (at the high end) against vocabulary and a less important function (in terms of accounting for between group variability) in which quantitative skills (an aptitude measure) is weighted positively and spatial ability negatively. The centroids for the seven groups are depicted in Figure 1. The unsuccessful group did rather poorer quantitatively than verbally. With the relatively large sample sizes it is not surprising that the mean vectors for the seven groups were found by a x^2 test to be significantly different.

The predictor selections and regression weights obtained from the weight development samples are summarized in Table 1. There is some variability in choice of predictors from sub group to sub group but, as expected, the high school grade point averages bulk large in these predictions. The variability in level of predictability squares fairly well with what we have observed earlier with respect to the predictability of earned grades in specific areas. Performance in biological science courses is relatively more predictable and performance in business courses relatively less predictable. For the successful students these specialized courses could be expected to bulk large in their cumulative performances.

Finally, the results of central interest to this study. The average squared error of prediction (discrepancy between predicted and earned cumulative GPA) in the validational sample using the total group prediction equation, the final one in Table 1, was .345. When predictions were made

Insert Table 1 and Figure 1 about here

for this same validational sample using individually adjusted regression weights, forming a linear combination of the entries in the first seven columns of Table 1 dependent upon the estimated probabilities of group membership, the overall average squared error of prediction was a bit smaller, .329.

Discussion

The results just stated for the validation sample might be considered positive only by one as accustomed as the present investigator to resounding negative results whenever patterns, profiles, interactions, or moderators as contributors to prediction have been sought. To find no loss of accuracy at the time of validation when antecedent measures are allowed to contribute nonlinearly to prediction was pleasantly surprising. The more so as, in this admittedly rough trial, there was little reason to believe in advance that either the classification groups or antecedent measures would be at all well-suited to such an approach. The subgrouping of students into conglomerates based on major at graduation can hardly be defended as likely to yield statistically homogeneous and nicely discriminable classes. The "unsuccessful" group, additionally, was an uncomfortable complement to these groups and the union of these probably corresponds only very roughly to that naturally occurring population, the entering freshman class. The antecedent measures were utilized because they were conveniently available. Though they have a well-established utility as linear contributors to the prediction of academic performance there was only limited empirical evidence that they could be used to discriminate academic group membership (Lunneborg and Lunneborg, 1970) and no suggestion that they possessed any nonlinear power.

Although the non-negative results obtained here are not strong enough to justify changing the mode of prediction for the prediction task illustrated here it may be worth speculating as to why the nonlinear treatment of the data did not break down on validation, as is so common. What may have been important was the de-coupling of the discrimination problem from the prediction problem. Earlier attempts by this investigator at pattern prediction treated the predictor or antecedent measure pool symmetrically and looked for patterns among the predictors rather than sought for any partitioning between linear predictors and other variables which can serve to pattern or moderate the contributions of the first.

If a technique as the one illustrated has promise it would seem to direct our attention, independently of the classic predictor selection question, to the search for coherent criterion subgroups and for measures which will reliably discriminate among them. Indeed, it could well be profitable to investigate techniques such as latent class analysis (Green, 1951) as part of a three stage development of a prediction system. The first stage would involve definition of optimal subclasses, the second the identification and weighting of variables to discriminate among these subclasses, and the third the moderating of classically determined linear predictions by these class discriminating measures, perhaps in the probabilistic sense illustrated here. In any applied setting the first two stages would have to be iterative, working back and forth between the definition of subclasses and the search for discriminating measures until classes that are optimally discriminable are obtained. For better or worse this investigator is intrigued enough to look again at nonlinear prediction.

References

Cooley, W. W., and Lohnes, P. R. Multivariate data analysis. New York: Wiley, 1971.

Dixon, W. J. (Ed.) BMD: Biomedical computer programs. Berkeley: University of California, 1968.

Efroymson, M. A. Multiple regression analysis. In A. Ralston and H. S. Wilf (Eds.), Mathematical methods for digital computers. New York: Wiley, 1960.

Green, B. F. A general solution for the latent class model of latent structure analysis. Psychometrika, 1951, 16, 151-166.

Lunneborg, C. E., & Lunneborg, P. W. Forecasting university major with the Washington pre-college test using discriminant functions. Seattle: Bureau of Testing, University of Washington, 1970. (Duplicated report)

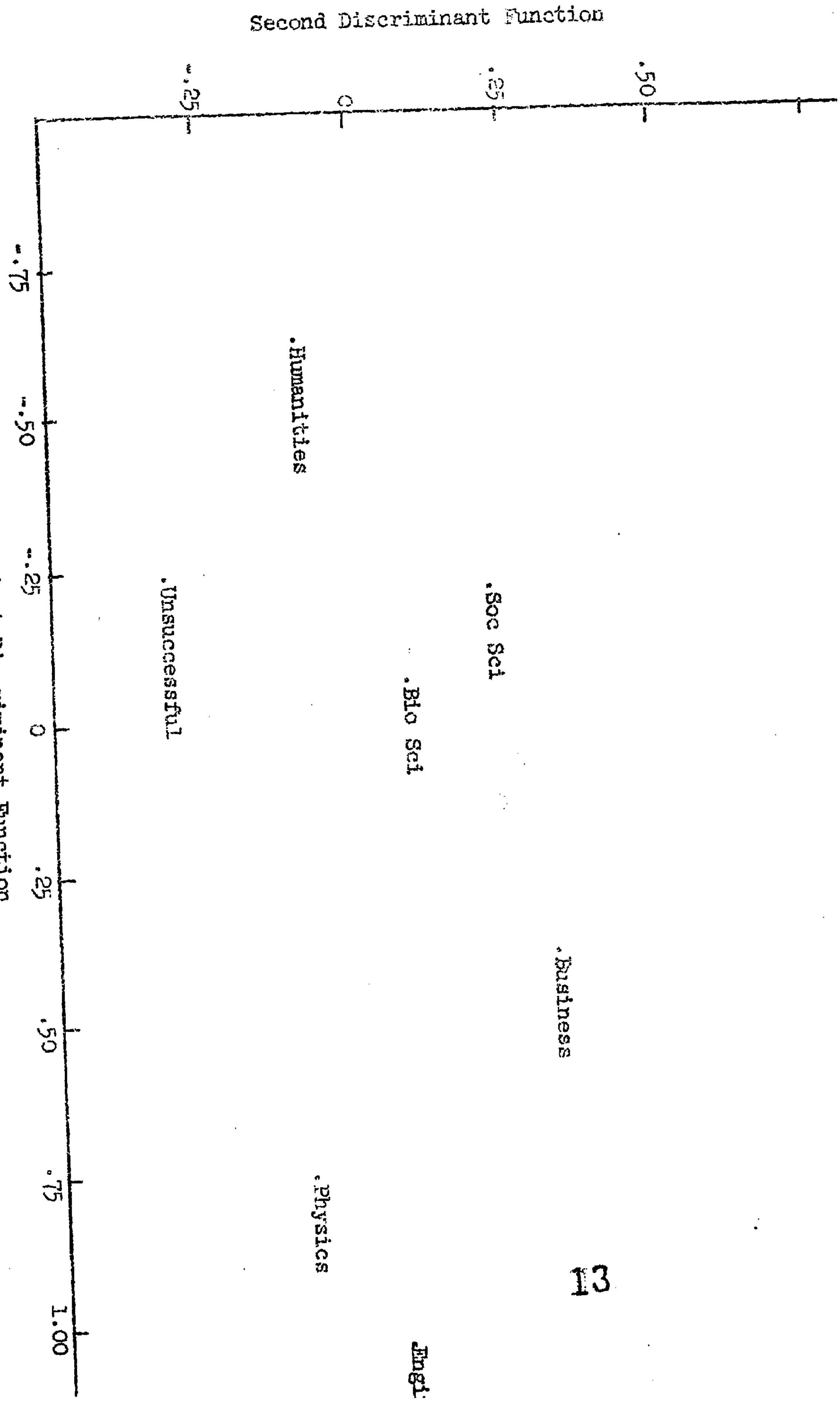
Second Discriminant Function
-.25
0
.25
.50
-.75
-.50
-.25
0
.25
.50
.75
1.00
.Humanities
.Unsuccessful
.Soc Sci
.Bio Sci
.Business
.Physics
Engi

Table 1

Predictor Selections, Standard Partial Regression Weights and Weight Sample Multiple Correlations

	Humanities (n = 142)	Physical Science (n = 97)	Social Science (n = 139)	Business (n = 61	Biological Science (n = 14)	Engineering (n = 114)	Unsuccessful (n = 194)	Total (n = 890)
English Usage	.138		.125					.050
Spelling		.075	-.038			-.010		
Read Comp			.162	.126		.050	.085	.062
Applied Math		.288		.175	.286	.392		-.026
HS English GPA	.135			.150			.113	.105
HS For Lang GPA					.109			.036
HS Math GPA	.159	.196	.236		.144	.209	.121	.113
HS Nat Sci GPA		.299		.080	.215	.106	.101	.103
HS Soc Sci GPA	.218		.340	.295	.360		.214	.216
HS Elect GPA	-.038	-.068						.015
Math Achieve								.075
Mech Reas								
Vocabulary								.024
Quant Skills								.028
Spatial Ability								.053
R	.495	.563	.559	.481	.576	.567	.536	[illegible]